

---

# Principles of

---

# Optimal Design

---

## Modeling and Computation

---

**SECOND EDITION**

PANOS Y. PAPALAMBROS

*University of Michigan*

DOUGLASS J. WILDE

*Stanford University*



**CAMBRIDGE**  
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE  
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK <http://www.cup.cam.ac.uk>

40 West 20th Street, New York, NY 10011-4211, USA <http://www.cup.org>

10 Stamford Road, Oakleigh, Melbourne 3166, Australia

Ruiz de Alarcón 13, 28014 Madrid, Spain

© Cambridge University Press 2000

This book is in copyright. Subject to statutory exception  
and to the provisions of relevant collective licensing agreements,  
no reproduction of any part may take place without  
the written permission of Cambridge University Press.

First published 2000

Printed in the United States of America

*Typefaces* Times Roman 10.75/13.5 pt. and Univers *System* L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub> [TB]

*A catalog record for this book is available from the British Library.*

*Library of Congress Cataloging in Publication Data*

Papalambros, Panos Y.

Principles of optimal design : modeling and computation / Panos Y. Papalambros,

Douglass J. Wilde. – 2nd ed.

p. cm.

Includes bibliographical references.

ISBN 0-521-62215-8

1. Mathematical optimization. 2. Mathematical models. I. Wilde, Douglass J.

II. Title.

QA402.5.P374 2000

519.3 – dc21

99-047982

ISBN 0 521 62215 8 hardback

ISBN 0 521 62727 3 paperback

---

# Contents

---

<i>Preface to the Second Edition</i>	<i>page</i> xiii
<i>Notation</i>	xvii
<b>1 Optimization Models</b>	1
1.1 Mathematical Modeling	1
The System Concept • Hierarchical Levels • Mathematical Models • Elements of Models • Analysis and Design Models • Decision Making	
1.2 Design Optimization	10
The Optimal Design Concept • Formal Optimization Models • Multicriteria Models • Nature of Model Functions • The Question of Design Configuration • Systems and Components • Hierarchical System Decomposition	
1.3 Feasibility and Boundedness	23
Feasible Domain • Boundedness • Activity	
1.4 Topography of the Design Space	30
Interior and Boundary Optima • Local and Global Optima • Constraint Interaction	
1.5 Modeling and Computation	38
1.6 Design Projects	39
1.7 Summary	39
Notes Exercises	
<b>2 Model Construction</b>	44
2.1 Modeling Data	44
Graphical and Tabular Data • Families of Curves • Numerically Generated Data	
2.2 Best Fit Curves and Least Squares	49

2.3	Neural Networks	51
2.4	Kriging	54
2.5	Modeling a Drive Screw Linear Actuator	57
	Assembling the Model Functions • Model Assumptions • Model Parameters • Negative Null Form	
2.6	Modeling an Internal Combustion Engine	62
	Flat Head Chamber Design • Compound Valve Head Chamber Design	
2.7	Design of a Geartrain	71
	Model Development • Model Summary • Model Reduction	
2.8	Modeling Considerations Prior to Computation	79
	Natural and Practical Constraints • Asymptotic Substitution • Feasible Domain Reduction	
2.9	Summary	83
	Notes Exercises	
<b>3</b>	<b>Model Boundedness</b>	87
3.1	Bounds, Extrema, and Optima	87
	Well-Bounded Functions • Nonminimizing Lower Bound • Multivariable Extension • Air Tank Design	
3.2	Constrained Optimum	92
	Partial Minimization • Constraint Activity • Cases	
3.3	Underconstrained Models	98
	Monotonicity • First Monotonicity Principle • Criticality • Optimizing a Variable Out • Adding Constraints	
3.4	Recognizing Monotonicity	103
	Simple and Composite Functions • Integrals	
3.5	Inequalities	105
	Conditional Criticality • Multiple Criticality • Dominance • Relaxation • Uncriticality	
3.6	Equality Constraints	109
	Equality and Activity • Replacing Monotonic Equalities by Inequalities • Directing an Equality • Regional Monotonicity of Nonmonotonic Constraints	
3.7	Variables Not in the Objective	114
	Hydraulic Cylinder Design • A Monotonicity Principle for Nonobjective Variables	
3.8	Nonmonotonic Functions	116
3.9	Model Preparation Procedure	119
3.10	Summary	121
	Notes Exercises	

---

<b>4 Interior Optima</b>	128
4.1 Existence	129
The Weierstrass Theorem • Sufficiency	
4.2 Local Approximation	131
Taylor Series • Quadratic Functions • Vector Functions	
4.3 Optimality	137
First-Order Necessity • Second-Order Sufficiency • Nature of Stationary Points	
4.4 Convexity	143
Convex Sets and Functions • Differentiable Functions	
4.5 Local Exploration	149
Gradient Descent • Newton’s Method	
4.6 Searching along a Line	154
Gradient Method • Modified Newton’s Method	
4.7 Stabilization	157
Modified Cholesky Factorization	
4.8 Trust Regions	160
Moving with Trust • Trust Region Algorithm	
4.9 Summary	163
Notes Exercises	
<b>5 Boundary Optima</b>	168
5.1 Feasible Directions	168
5.2 Describing the Constraint Surface	171
Regularity • Tangent and Normal Hyperplanes	
5.3 Equality Constraints	174
Reduced (Constrained) Gradient • Lagrange Multipliers	
5.4 Curvature at the Boundary	180
Constrained Hessian • Second-Order Sufficiency • Bordered Hessians	
5.5 Feasible Iterations	186
Generalized Reduced Gradient Method • Gradient Projection Method	
5.6 Inequality Constraints	194
Karush–Kuhn–Tucker Conditions • Lagrangian Standard Forms	
5.7 Geometry of Boundary Optima	198
Interpretation of KKT Conditions • Interpretation of Sufficiency Conditions	
5.8 Linear Programming	203
Optimality Conditions • Basic LP Algorithm	
5.9 Sensitivity	214
Sensitivity Coefficients	

5.10	Summary	216
	Notes Exercises	
<b>6</b>	<b>Parametric and Discrete Optima</b>	<b>223</b>
6.1	Parametric Solution	224
	Particular Optimum and Parametric Procedures • Branching • Graphical Interpretation • Parametric Tests	
6.2	The Monotonicity Table	232
	Setting up • First New Table: Reduction • Second New Table: Two Directions and Reductions • Third New Table: Final Reduction • Branching by Conditional Criticality • The Stress-Bound Cases • Parametric Optimization Procedure	
6.3	Functional Monotonicity Analysis	240
	Explicit Algebraic Elimination • Implicit Numerical Solution • Optimization Using Finite Element Analysis	
6.4	Discrete Variables	245
6.5	Discrete Design Activity and Optimality	247
	Constraint Activity Extended • Discrete Local Optima	
6.6	Transformer Design	255
	Model Development • Preliminary Set Constraint Tightening	
6.7	Constraint Derivation	259
	Discriminant Constraints • Constraint Addition • Linear and Hyperbolic Constraints • Further Upper and Lower Bound Generation • Case Analysis • Constraint Substitution: Remaining Cases	
6.8	Relaxation and Exhaustive Enumeration	270
	Continuous Relaxation: Global Lower Bounds • Problem Completion: Exhaustive Enumeration	
6.9	Summary	272
	Notes Exercises	
<b>7</b>	<b>Local Computation</b>	<b>278</b>
7.1	Numerical Algorithms	279
	Local and Global Convergence • Termination Criteria	
7.2	Single Variable Minimization	285
	Bracketing, Sectioning, and Interpolation • The Davies, Swann, and Campey Method • Inexact Line Search	
7.3	Quasi-Newton Methods	296
	Hessian Matrix Updates • The DFP and BFGS Formulas	
7.4	Active Set Strategies	300
	Adding and Deleting Constraints • Lagrange Multiplier Estimates	
7.5	Moving along the Boundary	305

---

7.6	Penalties and Barriers	306
	Barrier Functions • Penalty Functions • Augmented Lagrangian (Multiplier) Methods	
7.7	Sequential Quadratic Programming	313
	The Lagrange–Newton Equations • Enhancements of the Basic Algorithm • Solving the Quadratic Subproblem	
7.8	Trust Regions with Constraints	320
	Relaxing Constraints • Using Exact Penalty Functions • Modifying the Trust Region and Accepting Steps • Yuan’s Trust Region Algorithm	
7.9	Convex Approximation Algorithms	324
	Convex Linearization • Moving Asymptotes • Choosing Moving Asymptotes and Move Limits	
7.10	Summary	329
	Notes Exercises	
<b>8</b>	<b>Principles and Practice</b>	337
8.1	Preparing Models for Numerical Computation	338
	Modeling the Constraint Set • Modeling the Functions • Modeling the Objective	
8.2	Computing Derivatives	342
	Finite Differences • Automatic Differentiation	
8.3	Scaling	348
8.4	Interpreting Numerical Results	352
	Code Output Data • Degeneracy	
8.5	Selecting Algorithms and Software	354
	Partial List of Software Packages • Partial List of Internet Sites	
8.6	Optimization Checklist	358
	Problem Identification • Initial Problem Statement • Analysis Models • Optimal Design Model • Model Transformation • Local Iterative Techniques • Final Review	
8.7	Concepts and Principles	362
	Model Building • Model Analysis • Local Searching	
8.8	Summary	366
	Notes	
	<i>References</i>	369
	<i>Author Index</i>	381
	<i>Subject Index</i>	385

## Optimization Models

For the goal is not the last, but the best.  
*Aristotle (Second Book of Physics) (384–322 B.C.)*

Designing is a complex human process that has resisted comprehensive description and understanding. All artifacts surrounding us are the results of designing. Creating these artifacts involves making a great many decisions, which suggests that designing can be viewed as a *decision-making process*. In the decision-making paradigm of the design process we examine the intended artifact in order to identify possible alternatives and select the most suitable one. An abstract description of the artifact using mathematical expressions of relevant natural laws, experience, and geometry is the *mathematical model* of the artifact. This mathematical model may contain many alternative designs, and so criteria for comparing these alternatives must be introduced in the model. Within the limitations of such a model, the best, or *optimum*, design can be identified with the aid of mathematical methods.

In this first chapter we define the design optimization problem and describe most of the properties and issues that occupy the rest of the book. We outline the limitations of our approach and caution that an “optimum” design should be perceived as such only within the scope of the mathematical model describing it and the inevitable subjective judgment of the modeler.

### 1.1 Mathematical Modeling

Although this book is concerned with *design*, almost all the concepts and results described can be generalized by replacing the word *design* by the word *system*. We will then start with discussing mathematical models for general systems.

#### The System Concept

A system may be defined as a collection of entities that perform a specified set of tasks. For example, an automobile is a system that transports passengers. It follows that a system performs a *function*, or process, which results in an *output*. It is implicit that a system operates under causality, that is, the specified set of tasks is performed because of some stimulation, or *input*. A *block diagram*, Figure 1.1, is

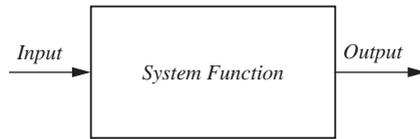


Figure 1.1. Block diagram representation.

a simple representation of these system elements. Causality generally implies that a dynamic behavior is possible. Thus, inputs to a system are entities identified to have an observable effect on the behavior of the system, while outputs are entities measuring the response of the system.

Although inputs are clearly part of the system characterization, what exactly constitutes an input or output depends on the *viewpoint* from which one observes the system. For example, an automobile can be viewed differently by an automaker's manager, a union member, or a consumer, as in Figure 1.2. A real system remains the same no matter which way you look at it. However, as we will see soon, the definition of a system is undertaken for the purpose of analysis and understanding; therefore the goals of this undertaking will influence the way a system is viewed. This may appear a trivial point, but very often it is a major block in communication between individuals coming from different backgrounds or disciplines, or simply having different goals.

### Hierarchical Levels

To study an object effectively, we always try to isolate it from its environment. For example, if we want to apply elasticity theory on a part to determine stresses and deflections, we start by creating the *free-body diagram* of the part, where the points of interaction with the environment are substituted by equivalent forces and moments. Similarly, in a thermal process, if we want to apply the laws of mass and energy

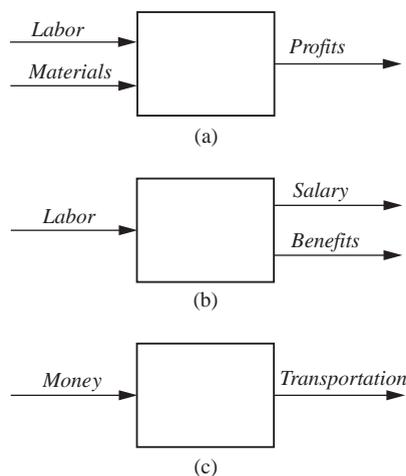


Figure 1.2. Viewpoints of system: automobile. (a) Manufacturer manager; (b) union member; (c) consumer.

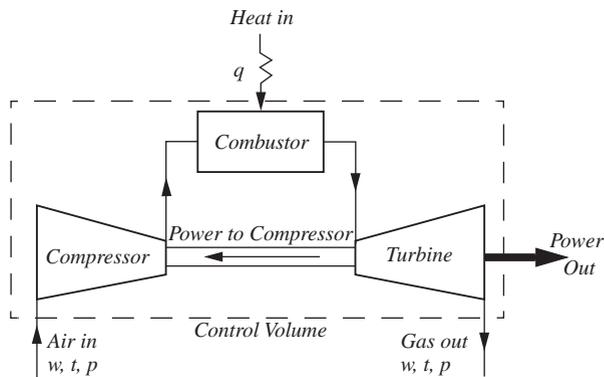


Figure 1.3. A gas-turbine system.

conservation to determine flow rates and temperatures, we start by specifying the *control volume*. Both the control volume and the free-body diagram are descriptions of the *system boundary*. Anything that “crosses” this boundary is a link between the system and its environment and will represent an input or an output characterizing the system.

As an example, consider the nonregenerative gas-turbine cycle in Figure 1.3. Drawing a control volume, we see that the links with the environment are the intake of the compressor, the exhaust of the turbine, the fuel intake at the combustor, and the power output at the turbine shaft. Thus, the air input (mass flow rate, temperature, pressure) and the heat flow rate can be taken as the inputs to the system, while the gas exit (mass flow rate, temperature, pressure) and the power takeoff are the outputs of the system. A simple block diagram would serve. Yet it is clear that the box in the figure indeed contains the components: compressor, combustor, turbine, all of which are themselves complicated machines. We see that the original system is made up of components that are systems with their own functions and input/output characterization. Furthermore, we can think of the gas-turbine plant as actually a component of a combined gas- and steam-turbine plant for liquefied petroleum. The original system has now become a component of a larger system.

The above example illustrates an important aspect of a system study: *Every system is analyzed at a particular level of complexity that corresponds to the interests of the individual who studies the system.* Thus, we can identify *hierarchical levels* in the system definition. Each system is broken down into subsystems that can be further broken down, with the various subsystems or components being interconnected. A boundary around any subsystem will “cut across” the links with its environment and determine the input/output characterization. These observations are very important for an appropriate identification of the system that will form the basis for constructing a mathematical model.

We may then choose to represent a system as a single unit at one level or as a collection of subsystems (for example, components and subcomponents) that must be coordinated at an overall “system level.” This is an important modeling decision when the size of the system becomes large.

### Mathematical Models

A real system, placed in its real environment, represents a very complex situation. The scientist or the engineer who wishes to study a real system must make many concessions to reality to perform some analysis on the system. It is safe to say that in practice *we never analyze a real system but only an abstraction of it*. This is perhaps the most fundamental idea in engineering science and it leads to the concept of a model:

*A model is an abstract description of the real world giving an approximate representation of more complex functions of physical systems.*

The above definition is very general and applies to many different types of models. In engineering we often identify two broad categories of models: *physical* and *symbolic*. In a physical model the system representation is a tangible, material one. For example, a scale model or a laboratory prototype of a machine would be a physical model. In a symbolic model the system representation is achieved by means of all the tools that humans have developed for abstraction—drawings, verbalization, logic, and mathematics. For example, a machine blueprint is a *pictorial* symbolic model. Words in language are models and not the things themselves, so that when they are connected with logical statements they form more complex *verbal* symbolic models. Indeed, the artificial computer languages are an extension of these ideas.

The symbolic model of interest here is the one using a *mathematical* description of reality. There are many ways that such models are defined, but following our previous general definition of a model we can state that:

*A mathematical model is a model that represents a system by mathematical relations.*

The simplest way to illustrate this idea is to look back at the block diagram representation of a system shown in Figure 1.1. Suppose that the output of the system is represented by a quantity  $y$ , the input by a quantity  $x$ , and the system function by a *mathematical function*  $f$ , which calculates a value of  $y$  for each value of  $x$ . Then we can write

$$y = f(x). \tag{1.1}$$

This equation is the mathematical model of the system represented in Figure 1.1. From now on, when we refer to a model we imply a mathematical one.

The creation of modern science follows essentially the same path as the creation of mathematical models representing our world. Since by definition a model is only an approximate description of reality, we anticipate that there is a varying degree of success in model construction and/or usefulness. A model that is successful and is supported by accumulated empirical evidence often becomes a *law* of science.

*Virtual reality* models are increasingly faithful representations of physical systems that use computations based on mathematical models, as opposed to realistic-looking effects in older computer games.

### Elements of Models

Let us consider the gas-turbine example of Figure 1.3. The input air for the compressor may come directly from the atmosphere, and so its temperature and pressure will be in principle beyond the power of the designer (unless the design is changed or the plant is moved to another location). The same is true for the output pressure from the turbine, since it exhausts in the atmosphere. The unit may be specified to produce a certain amount of net power. The designer takes these as given and tries to determine required flow rates for air and fuel, intermediate temperatures and pressures, and feedback power to the compressor. To model the system, the laws of thermodynamics and various physical properties must be employed. Let us generalize the situation and identify the following model elements for all systems:

*System Variables.* These are quantities that specify different states of a system by assuming different values (possibly within acceptable ranges). In the example above, some variables can be the airflow rate in the compressor, the pressure out of the compressor, and the heat transfer rate into the combustor.

*System Parameters.* These are quantities that are given *one* specific value in any particular model statement. They are fixed by the application of the model rather than by the underlying phenomenon. In the example, atmospheric pressure and temperature and required net power output will be parameters.

*System Constants.* These are quantities fixed by the underlying phenomenon rather than by the particular model statement. Typically, they are natural constants, for example, a gas constant, and the designer cannot possibly influence them.

*Mathematical Relations.* These are equalities and inequalities that relate the system variables, parameters, and constants. The relations include some type of functional representation such as Equation (1.1). Stating these relations is the most difficult part of modeling and often such a relation is referred to as *the* model. These relations attempt to describe the function of the system within the conditions imposed by its environment.

The clear distinction between variables and parameters is very important at the modeling stage. The choice of what quantities will be classified as variables or parameters is a *subjective* decision dictated by choices in hierarchical level, boundary isolation, and intended use of the model of the system. This issue is addressed on several occasions throughout the book.

As a final note, it should be emphasized that the mathematical representation  $y = f(x)$  of the system function is more symbolic than real. The actual “function” may be a system of equations, algebraic or differential, or a computer-based procedure or subroutine.

### Analysis and Design Models

Models are developed to increase our understanding of how a system works. A *design* is also a system, typically defined by its geometric configuration, the materials used, and the task it performs. To model a design mathematically we must be able to define it completely by assigning values to each quantity involved, with these values satisfying mathematical relations representing the performance of a task.

In the traditional approach to design it has been customary to distinguish between *design analysis* and *design synthesis*. Modeling for design can be thought of in a similar way. In the model description we have the same elements as in general system models: design variables, parameters, and constants. To determine how these quantities relate to each other for proper performance of function of the design, we must first conduct *analysis*. Examples can be free-body diagram analysis, stress analysis, vibration analysis, thermal analysis, and so on. Each of these analyses represents a descriptive model of the design. If we want to predict the overall performance of the design, we must construct a model that incorporates the results of the analyses. Yet its goals are different, since it is a predictive model. Thus, in a design modeling study we must distinguish between *analysis models* and *design models*. Analysis models are developed based on the principles of engineering science, whereas design models are constructed from the analysis models for specific prediction tasks and are problem dependent.

As an illustration, consider the straight beam formula for calculating bending stresses:

$$\sigma = My/I, \tag{1.2}$$

where  $\sigma$  is the normal stress at a distance  $y$  from the neutral axis at a given cross section,  $M$  is the bending moment at that cross section, and  $I$  is the moment of inertia of the cross section. Note that Equation (1.2) is valid only if several simplifying assumptions are satisfied. Let us apply this equation to the trunk of a tree subjected to a wind force  $F$  at a height  $h$  above the ground (Alexander 1971), as in Figure 1.4(a). If the tree has a circular trunk of radius  $r$ , the moment of inertia is  $I = \pi r^4/4$  and the maximum bending stress is at  $y = r$ :

$$\sigma_{\max} = 4Fh/\pi r^3. \tag{1.3}$$

If we take the tree as given (i.e.,  $\sigma_{\max}$ ,  $h$ ,  $r$  are parameters), then Equation (1.3) solved for  $F$  can tell us the maximum wind force the tree can withstand before it breaks. Thus Equation (1.3) serves as an analysis model. However, a horticulturist may view this as a design problem and try to protect the tree from high winds by appropriately trimming the foliage to decrease  $F$  and  $h$ . Note that the force  $F$  would depend both on the wind velocity and the configuration of the foliage. Now Equation (1.3) is a design model with  $h$  and (partially)  $F$  as variables. Yet another situation exists in Figure 1.4(b) where the cantilever beam must be designed to carry the load  $F$ . Here the load is a parameter; the length  $h$  is possibly a parameter but the radius  $r$  would be normally considered as the design variable. The analysis model yields yet another design model.

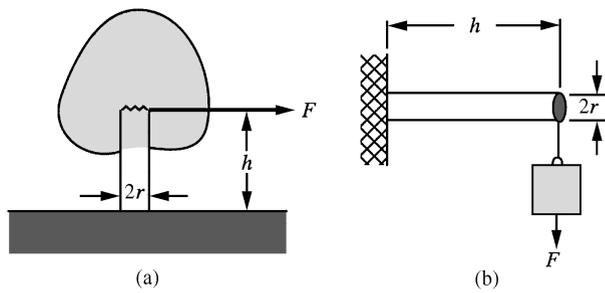


Figure 1.4. (a) Wind force acting on a tree trunk. (b) Cantilever beam carrying a load.

The analysis and design models may not be related in as simple a manner as above. If the analysis model is represented by a differential equation, the constants in this equation are usually design variables. For example, a gear motor function may be modeled by the equation of motion

$$J(d^2\theta/dt^2) + b(d\theta/dt) = -f_g r, \quad (1.4)$$

where  $J$  is the moment of inertia of the armature and pinion,  $b$  is the damping coefficient,  $f_g$  is the tangential gear force,  $r$  is the gear radius,  $\theta$  is the angle of rotation, and  $t$  is time. Here  $J$ ,  $b$ , and  $f_g r$  are constants for the differential equation.

However, the design problem may be to determine proper values for gear and shaft sizes, or the natural frequency of the system, which would require making  $J$ ,  $b$ , and  $r$  design variables. An explicit relation among these variables would require solving the differential equation each time with different (numerical) values for its constants. If the equation cannot be solved explicitly, the design model would be represented by a computer subroutine that solves the equation iteratively.

Before we conclude this discussion we must stress that there is no single design model, but different models are constructed for different needs. The analysis models are much more restricted in that sense, and, once certain assumptions have been made, the analysis model is usually unique. The importance of the influence of a given viewpoint on the design model is seen by another simple example. Let us examine a simple round shaft supported by two bearings and carrying a gear or pulley, as in Figure 1.5. If we neglect the change of diameters at the steps, we can say that the design of the shaft requires a choice of the diameter  $d$  and a material with associated properties such as density, yield strength, ultimate strength, modulus of elasticity, and fatigue endurance limit. Because the housing is already specified, the length between the supporting bearings,  $l$ , cannot be changed. Furthermore, suppose that we have in stock only one kind of steel in the diameter range we expect.

Faced with this situation, the diameter  $d$  will be the only design variable we can use; the material properties and the length  $l$  would be considered as design parameters. This is what the viewpoint of the shaft designer would be. However, suppose that after some discussion with the housing designer, it is decided that changes in the housing dimensions might be possible. Then  $l$  could be made a variable. The project manager,

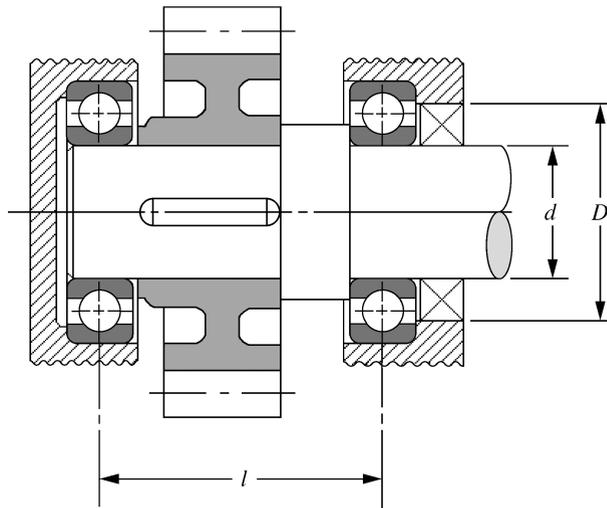


Figure 1.5. Sketch of a shaft design.

who might order any materials *and* change the housing dimensions, would view  $d$ ,  $l$ , and material properties all as design variables. In each of the three cases, the model will be different and of course this would also affect the results obtained from it.

### Decision Making

We pointed out already that design models are predictive in nature. This comes rather obviously from our desire to study how a design performs and how we can influence its performance. The implication then is that a design can be modified to generate different alternatives, and the purpose of a study would be to select “the most desirable” alternative. Once we have more than one alternative, a need arises for making a decision and choosing one of them. Rational choice requires a *criterion* by which we evaluate the different alternatives and place them in some form of ranking. This criterion is a new element in our discussion on design models, but in fact it is always implicitly used any time a design is selected.

A criterion for evaluating alternatives and choosing the “best” one cannot be unique. Its choice will be influenced by many factors such as the design application, timing, point of view, and judgment of the designer, as well as the individual’s position in the hierarchy of the organization. To illustrate this, let us return to the shaft design example. One possible criterion is lightweight construction so that weight can be used to generate a ranking, the “best” design being the one with minimum weight. Another criterion could be rigidity, so that the design selected would have maximum rigidity for, say, best meshing of the attached gears. For the shop manager the ease of manufacturing would be more important so that the criterion then would be the sum of material and manufacturing costs. For the project or plant manager, a minimum cost design would be again the criterion but now the shaft cost would not be examined alone, but in conjunction with the costs of the other parts that the

shaft has to function with. A corporate officer might add possible liability costs and so on.

A criterion may change with time. An example is the U.S. automobile design where best performance measures shifted from maximum power and comfort to maximum fuel economy and more recently to a rather unclear combination of criteria for maximum quality and competitiveness. One may argue that the ultimate criterion is always cost. But it is not always practical to use cost as a criterion because it can be very difficult to quantify. Thus, the criterion quantity shares the same property as the other elements of a model: It is an approximation to reality and is useful within the limitations of the model assumptions.

A design model that includes an evaluation criterion is a *decision-making model*. More often this is called an *optimization model*, where the “best” design selected is called the optimal design and the criterion used is called the *objective* of the model. We will study some optimization models later, but now we want to discuss briefly the ways design optimization models can be used in practice.

The motivation for using design optimization models is the selection of a good design representing a compromise of many different requirements with little or no aid from prototype hardware. Clearly, if this attempt is successful, substantial cost and design cycle time savings will be realized. Such optimization studies may provide the competitive edge in product design.

In the case of *product development*, a new original design may be represented by its model. Before any hardware are produced, design alternatives can be generated by manipulating the values of the design variables. Also, changes in design parameters can show the effect of external factor changes on a particular design. The objective criterion will help select the best of all generated alternatives. Consequently, a preliminary design is developed. How good it is depends on the model used. Many details must be left out because of modeling difficulties. But with accumulated experience, reliable elaborate models can be constructed and *design costs* will be drastically reduced. Moreover, the construction, validation, and implementation of a design model in the computer may take very much less time than prototype construction, and, when a prototype is eventually constructed, it will be much closer to the desired production configuration. Thus, design cycle time may be also drastically reduced.

In the case of *product enhancement*, an existing design can be described by a model. We may not be interested in drastic design changes that might result from a full-scale optimization study but in relatively small design changes that might improve the performance of the product. In such circumstances, the model can be used to predict the effect of the changes. As before, design cost and cycle time will be reduced. Sometimes this type of model use is called a *sensitivity study*, to be distinguished from a complete *optimization study*.

An optimization study usually requires several iterations performed in the computer. For large, complicated systems such iterations may be expensive or take too much time. Also, it is possible that a mathematical optimum could be difficult to locate precisely. In these situations, a complete optimization study is not performed.

Instead, several iterations are made until a sufficient improvement in the design has been obtained. This approach is often employed by the aerospace industry in the design of airborne structures. A design optimization model will use structural (typically finite element) and fluid dynamics analysis models to evaluate structural and aerodynamic performance. Every design iteration will need new analyses for the values of the design variables at the current iteration. The whole process becomes very demanding when the level of design detail increases and the number of variables becomes a few hundred. Thus, the usual practice is to stop the iterations when a competitive weight reduction is achieved.

## 1.2 Design Optimization

### The Optimal Design Concept

The concept of design was born the first time an individual created an object to serve human needs. Today design is still the ultimate expression of the art and science of engineering. From the early days of engineering, the goal has been *to improve the design so as to achieve the best way of satisfying the original need, within the available means.*

The design process can be described in many ways, but we can see immediately that there are certain elements in the process that any description must contain: a *recognition of need*, an *act of creation*, and a *selection of alternatives*. Traditionally, the selection of the “best” alternative is the phase of *design optimization*. In a traditional description of the design phases, recognition of the original need is followed by a technical statement of the problem (problem definition), the creation of one or more physical configurations (synthesis), the study of the configuration’s performance using engineering science (analysis), and the selection of “best” alternative (optimization). The process concludes with testing of the prototype against the original need.

Such sequential description, though perhaps useful for educational purposes, cannot describe reality adequately since the question of how a “best” design is selected within the available means is pervasive, influencing all phases where decisions are made.

So what is design optimization?

We defined it loosely as the selection of the “best” design within the available means. This may be intuitively satisfying; however, both to avoid ambiguity and to have an operationally useful definition we ought to make our understanding rigorous and, ideally, quantifiable. We may recognize that a rigorous definition of “design optimization” can be reached if we answer the questions:

1. How do we describe different designs?
2. What is our criterion for “best” design?
3. What are the “available means”?

The first question was addressed in the previous discussion on design models, where a design was described as a system defined by design variables, parameters, and constants. The second question was also addressed in the previous section in the discussion on decision-making models where the idea of “best” design was introduced and the criterion for an optimal design was called an *objective*. The objective function is sometimes called a “cost” function since minimum cost often is taken to characterize the “best” design. In general, the criterion for selection of the optimal design is a function of the design variables in the model.

We are left with the last question on the “available means.” Living, working, and designing in a finite world obviously imposes limitations on what we may achieve. Brushing aside philosophical arguments, we recognize that any design decision will be subjected to limitations imposed by the natural laws, availability of material properties, and geometric compatibility. On a more practical level, the usual engineering specifications imposed by the clients or the codes must be observed. Thus, by “available means” we signify a set of requirements that must be satisfied by any acceptable design. Once again we may observe that these design requirements may not be uniquely defined but are under the same limitations as the choice of problem objective and variables. In addition, the choices of design requirements that must be satisfied are very intimately related to the choice of objective function and design variables.

As an example, consider again the shaft design in Figure 1.5. If we choose minimum weight as objective and diameter  $d$  as the design variable, then possible specifications are the use of a particular material, the fixed length  $l$ , and the transmitted loads and revolutions. The design requirements we may impose are that the maximum stress should not exceed the material strength and perhaps that the maximum deflection should not surpass a limit imposed by the need for proper meshing of mounted gears. Depending on the kind of bearings used, a design requirement for the slope of the shaft deflection curve at the supporting ends may be necessary. Alternatively, we might choose to maximize rigidity, seeking to minimize the maximum deflection as an objective. Now the design requirements might change to include a limitation in the space  $D$  available for mounting, or even the maximum weight that we can tolerate in a “lightweight” construction. We resolve this issue by agreeing that *the design requirements to be used are relative to the overall problem definition and might be changed with the problem formulation*. The design requirements pertaining to the current problem definition we will call *design constraints*. We should note that design constraints include all relations among the design variables that must be satisfied for proper functioning of the design.

So what *is* design optimization?

Informally, but rigorously, we can say that design optimization involves:

1. The selection of a set of variables to describe the design alternatives.
2. The selection of an objective (criterion), expressed in terms of the design variables, which we seek to minimize or maximize.

3. The determination of a set of constraints, expressed in terms of the design variables, which must be satisfied by any acceptable design.
4. The determination of a set of values for the design variables, which minimize (or maximize) the objective, while satisfying all the constraints.

By now, one should be convinced that this definition of optimization suggests a philosophical and tactical approach during the design process. It is not a phase in the process but rather a pervasive viewpoint.

Philosophically, optimization formalizes what humans (and designers) have always done. Operationally, it can be used in design, in any situation where analysis is used, and is therefore subjected to the same limitations.

### Formal Optimization Models

Our discussion on the informal definition of design optimization suggests that first we must formulate the problem and then solve it. There may be some iteration between formulation and solution, but, in any case, any quantitative treatment must start with a mathematical representation. To do this formally, we assemble all the design variables  $x_1, x_2, \dots, x_n$  into a *vector*  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  belonging to a subset  $\mathcal{X}$  of the  $n$ -dimensional real space  $\mathfrak{R}^n$ , that is,  $\mathbf{x} \in \mathcal{X} \subseteq \mathfrak{R}^n$ . The choice of  $\mathfrak{R}^n$  is made because the vast majority of the design problems we are concerned with here have real variables. The set  $\mathcal{X}$  could represent certain *ranges* of real values or certain *types*, such as integer or standard values, which are very often used in design specifications.

Having previously insisted that the objective and constraints must be quantifiably expressed in terms of the design variables, we can now assert that the objective is a *function* of the design variables, that is,  $f(\mathbf{x})$ , and that the constraints are represented by *functional relations* among the design variables such as

$$h(\mathbf{x}) = 0 \quad \text{and} \quad g(\mathbf{x}) \leq 0. \quad (1.5)$$

Thus we talk about *equality* and *inequality* constraints given in the form of equal to zero and less than or equal to zero. For example, in our previous shaft design, suppose we used a hollow shaft with outer diameter  $d_o$ , inner diameter  $d_i$ , and thickness  $t$ . These quantities could be viewed as design variables satisfying the equality constraint

$$d_o = d_i + 2t, \quad (1.6)$$

which can be rewritten as

$$d_o - d_i - 2t = 0 \quad (1.7)$$

so that the constraint function is

$$h(d_o, d_i, t) = d_o - d_i - 2t. \quad (1.8)$$

We could also have an inequality constraint specifying that the maximum stress does not exceed the strength of the material, for example,

$$\sigma_{\max} \leq S, \quad (1.9)$$

where  $S$  is some properly defined strength (i.e., maximum allowable stress). However,  $\sigma_{\max}$  should be expressed in terms of  $d_o$ ,  $d_i$ , and  $t$ . If we neglect the effect of bending for simplicity, we can write

$$\sigma_{\max} = \tau_{\max} = M_t(d_o/2)/J, \quad (1.10)$$

where  $M_t$  is the torsional moment and  $J$  is the polar moment of inertia,

$$J = (\pi/32)(d_o^4 - d_i^4). \quad (1.11)$$

At this point we may view (1.10) and (1.11) as additional equality constraints with  $\sigma_{\max}$  and  $J$  being additional design variables. Note that  $M_t$  would be a design parameter. Thus, we can rewrite them as follows:

$$\begin{aligned} \sigma_{\max} - S &\leq 0, \\ \sigma_{\max} - M_t(d_o/2J) &= 0, \\ J - (\pi/32)(d_o^4 - d_i^4) &= 0, \end{aligned} \quad (1.12)$$

so that we have one inequality and two equality constraints corresponding to (1.9). We could also eliminate  $\sigma_{\max}$  and  $J$  and get

$$16M_t d_o / \pi (d_o^4 - d_i^4) - S \leq 0, \quad (1.13)$$

that is, just one inequality constraint. This implies that  $\sigma_{\max}$  and  $J$  were considered *intermediate variables* that with the formulation (1.13) will disappear from the model statement. The above operation from (1.12) to (1.13) is a *model transformation* and it must be always performed judiciously so that the problem resulting from the transformation is *equivalent* to the original one and usually easier to solve. A strict definition of equivalence is difficult. Normally, we simply mean that the solution set of the transformed model is the same as that of the original model.

On the issue of transformation we may observe that the functional constraint representation (1.5) is not necessarily unique. For example, the renderings (1.7) and (1.13) of Equations (1.6) and (1.9), respectively, could have been written as

$$(d_o - d_i)/2t - 1 = 0, \quad (1.14)$$

$$16M_t d_o - S\pi d_o^4 + S\pi d_i^4 \leq 0. \quad (1.15)$$

The functions at the left side of (1.7) and (1.14) as well as (1.13) and (1.15) are *not the same*. For example, the function  $h$  in (1.8) varies linearly with  $t$ , which is not the case in (1.14). Of course, both functions were arrived at through transformations of the original (1.6). If we are careful, we should arrive at equivalent forms; yet very often careless transformations may confuse the analysis by introducing extraneous information not really there, or by hiding additional information. This is particularly